

BIG-DATA APPROACH IN ABUNDANCE ESTIMATION OF NON-IDENTIFIABLE ANIMALS WITH CAMERA-TRAPS AT THE SPOTS OF ATTRACTION

E.E. Ivanko, Institute of Mathematics and Mechanics, Ural Branch of the Russian Academy of Sciences, Yekaterinburg, Russian Federation; Ural Federal University, Yekaterinburg, Russian Federation, evgeny.ivanko@gmail.com

Camera-traps is a relatively new but already popular instrument in the estimation of abundance of non-identifiable animals. Although camera-traps are convenient in application, there remain both theoretical complications (such as spatial autocorrelation or false negative problem) and practical difficulties, for example, laborious random sampling. In the article I propose an alternative way to bypass the mentioned problems. In the proposed approach, the raw video information collected from the camera-traps situated at the spots of natural attraction is turned into the frequency of visits, and the latter is transformed into the desired abundance estimate. The key for such a transformation is the application of the correction coefficients, computed for each particular observation environment using the Bayesian approach and the massive database (DB) of observations under various conditions. The proposed method is based on automated video-capturing at a moderate number of easy to reach spots, so in the long term many laborious census works may be conducted easier, cheaper and cause less disturbance for the wild life. Information post-processing is strictly formalized, which leaves little chance for subjective alterations. However, the method heavily relies on the volume and quality of the DB, which in its turn heavily relies on the efforts of the community. Although the construction of such DB could be rather difficult and controversial, it is much easier than the solution of the initial abundance estimation problem. Moreover, such a rich DB of visits might benefit not only censuses, but also many behavioral studies.

Keywords: abundance estimation; camera-traps; big-data; Bayes naive classifier.

Abbreviations

AS – attractive spot(s), natural or artificial places of attraction for the selected species;
DB – massive database of the observations of different species at different types of attractive spots under various conditions;
NR – nature reserve, a protected area of biological, geological or other special conservation interest.

Introduction

Estimation of abundance [1, 2] of different species is required for many types of ecological study as well as consists an important part of duties of each nature reserve (NR). It helps to find the species in need of special attention and to balance the recreational load, to analyze the dynamics of populations and to reveal important ecological interconnections. Each NR can widen this list easily.

The obvious way to estimate abundance is just to count all the individuals of the given species within the region of interest. However, the situation where such a direct method can be applied in practice is very rare.

All other techniques may be considered indirect, where the researcher estimates the number of individuals under some limited conditions (both in space and time) and extrapolates the result. For the purposes of this article, it is convenient to divide the indirect techniques into two large groups: 1) with and 2) without individual identification. The backbone of the first group is the capture-recapture approach [1–3] broadened by such identification methods as camera-trap estimation of individually identifiable animals [4–13] or microsatellite analysis [14–20] of the collected biomaterial.

The capture-recapture method is very popular and effective, however, not many species have natural individual colouring or shape marks (like tiger's stripes or deer's antlers) allowing easy and practical application of camera-traps. For non-identifiable animals, the application of the classical capture-recapture method (both in the "traps" and "camera-traps" variations) is connected with considerable marking efforts and stress for the animals. One of the possible identifying techniques for non-identifiable animals – DNA analysis – remains a relatively difficult and expensive procedure; in addition, the process of samples collection is laborious and difficult to facilitate.

The modern answer to these challenges is *spatial* capture-recapture [21–24]. The approach of this article resembles to some extent a degenerate case of spatial capture-recapture method without marking, except that the "activity centers", which are responsible for the spatial distribution, are created artificially and can be observed directly.

Summarizing, the techniques of this group perfectly fit the species with natural marks or well predictable spatial distributions; otherwise, the researcher has to resort to relatively expensive and difficult artificial marking.

In the second group of estimation techniques corresponding to non-identifiable (or rather not easily identifiable) animals, distance sampling methods [25–27] dominate. Without individual identification, the researcher has only one type of data – the frequency of the observations (either of the species itself or some kind of its traces), which should be transformed into an estimation of the density.

Camera-traps is a relatively new but already popular and promising instrument in the estimation of abundance of non-identifiable animals [28–32]. One of the largest issues connected with the application of this instrument is the demanding requirements to the number of the employed camera-traps. Either the researcher has to provide a considerable amount of cameras (which is expensive in both cost and labour) or he runs the risk of underestimation or even false zero.

Here, I propose the idea of a novel community driven method of the estimation of abundance with camera-traps. Theoretically, this method allows each particular investigator to rely on the efforts of many distant independent researchers and get the desirable estimation at low expenses. In the proposed approach highly attractive spots (below, I use the acronym AS for both singular and plural) like feeders or marking spots play the role of point transects, which allows to collect rich observation statistics and to cope with the false negative problem [33, 34]. Another benefit of the AS usage is the reduced labour demands for the data collection, since often AS are relatively easy accessible. Of course, the observations conducted at the AS will result in an overrated observation frequency, which is further corrected by the coefficients computed basing on the Bayesian approach (see Section 2) applied to the specially prepared observation data (see Section 4). For simplicity, below I consider abundance of solitary animals, however it is possible to proceed to social animals as soon as the average group size is known.

1. AS: Types, Positioning and the Corresponding Area

In this section I discuss the notion of AS and speculate on the area which connects abundance with density paving the way for the abundance estimation.

All kinds of spots that attract the species of interest may be taken as AS, provided that three important conditions are met:

1) AS are placed inside the research area in such a way that almost every animal of the studied species living in (or regularly appearing in) the research area is aware of and periodically visits at least one AS. It is a separate difficult problem to place the AS correctly, but if I take, for example, feeders, NR encounters the positioning problem with the same criterion independently of camera-trap observations and solves it in some way thus I leave this discussion beyond the scope of the article.

2) There are no “empty” AS – all the AS of the selected kind within the research area are equipped with the camera-traps so that any event “an animal visits an AS” is recorded. This condition is easy to fulfill if artificial AS such as feeders are used. Otherwise the experimenter should be confident that all the AS within the research area are discovered.

3) It is practically possible to use camera-traps to fix an event “animal is at AS” clearly and to measure the duration of this event. Technically, once a camera fixes an animal, it may turn off and recheck the animal’s presence periodically; thus we get a satisfactory estimate of the visit duration, while saving the battery. The event “animal at AS” may be defined differently depending on the type of AS and on the species under research, but, for a particular AS type and a species, this event should be clearly described and commonly accepted by the professional community. For example, an elk may be considered staying at a feeder if and only if the distance between the elk and the feeder does not exceed 1 meter (or some other value accepted by the experts).

I mentioned the notion of “research area” above and even stipulated the restricting conditions on it. Let me clarify how this research area may be formally defined. In the proposed approach, camera-traps are used to measure the abundance over some area, after which the results are extrapolated to the desired part of the NR territory. The key to such an extrapolation is density $D = N/S$, where N is the estimated number of animals and S is the “area of the survey.” While N is just a number resulting from the observation by the methods discussed in the next section, the value of S becomes ambiguous in case I consider animals in the wild. For example, I conduct observations in a forest without any fence and use just a few cameras (see Fig. 1 a–1 d); which area should be preferred as a denominator in the expression for D ? If there are more cameras (like in Fig. 1 e–1 h), then the corresponding area appears a bit clearer. An enormous number of cameras (Fig. 1 i) is needed to define the corresponding area explicitly.

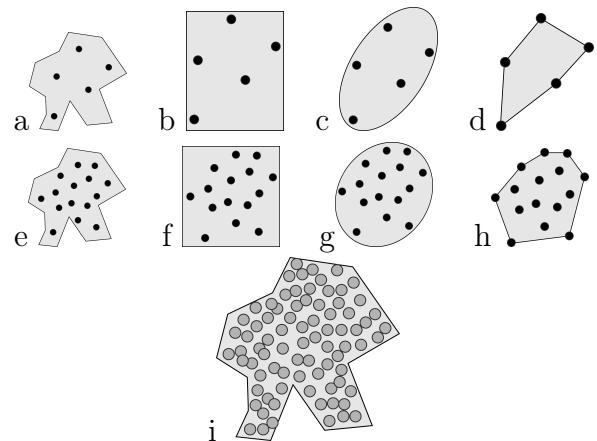


Fig. 1. Which area better corresponds to the AS positions?

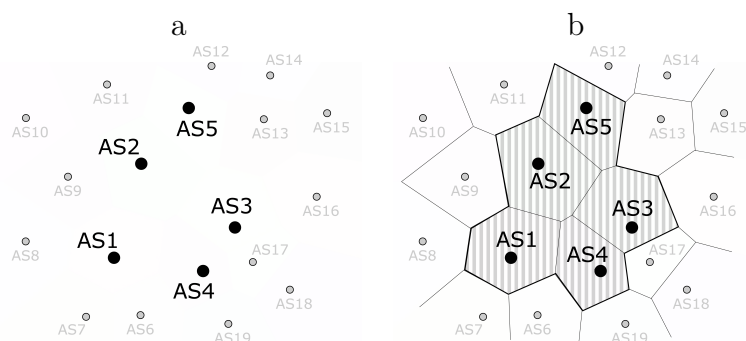


Fig. 2. The experiment area around the set of AS; a) a part of an infinite AS net around AS1–AS5; b) the area, corresponding to the set AS1–AS5 according to the Voronoi diagram

How can I rationally estimate the research area S ? Let me suppose for a moment that there is an infinite net of more or less regularly positioned cameras on the surface beyond finite set $C = \{AS1, \dots, AS5\}$ of the real AS used in the survey (see Fig. 2 a). In this case, the problem has a natural solution: I just need to surround the area “corresponding to” or “proportional to” the set C . Voronoi diagram [35,36] gives an elegant way to define the sought “corresponding” area formally (see Fig. 2 b).

In the observation survey, I do not have any other AS except for the real, so new virtual neighbours should be constructed artificially. It is possible to propose many ways to expand C to an infinite net with some regularity; let me describe one of them briefly.

Start with set C of real AS, used in the observation survey (see Fig. 3 a). At the first step, build the Delaunay triangulation [36,37] and calculate the average length R of its edges (See Fig. 3b). At the second step, construct the R -equidistant of set C (see Fig. 3 c, d). The final third step consists of the following short algorithm (see Fig. 3 e): 1) start from an arbitrary point on the equidistant (a better choice is an intersection of two neighbouring circles) $X1$; 2) construct the circle of radius R from center $X1$ and find $X2$ – the intersection of this circle with the equidistant in the clockwise direction from the starting point; 3) repeat the process from the newly constructed point until $X1$ is reached from another side.

2. Abundance Estimation

Suppose there is an unknown number N of animals and m stationary AS each of which is equipped with a camera-trap. The observation is conducted during time \mathbb{T} simultaneously for all the AS.

The conditions of the observation may change during the survey. The corresponding frequency and duration of visits as well as the resulted abundance estimate may vary significantly depending on the survey condition. For example, a part of an observation survey may pass under the rain, another part is conducted at night and so on. Let W be the number of different conditions during the observation survey. Each condition \mathcal{A}^j , where $j \in \overline{1, W}$, is described by the sequence of states $\mathcal{A}^j = (\alpha_1^j, \dots, \alpha_\Theta^j)$, lasts for the time $T(\mathcal{A}^j) \in [0, \mathbb{T}]$ and results in the estimated number of animals N_j .

For example, if a survey is conducted during both day and night and faces both dry and rainy weather in each time of day, then there are $W = 4$ different conditions: $\mathcal{A}^1 = (rain, day)$, $\mathcal{A}^2 = (rain, night)$, $\mathcal{A}^3 = (dry, day)$, $\mathcal{A}^4 = (dry, night)$.

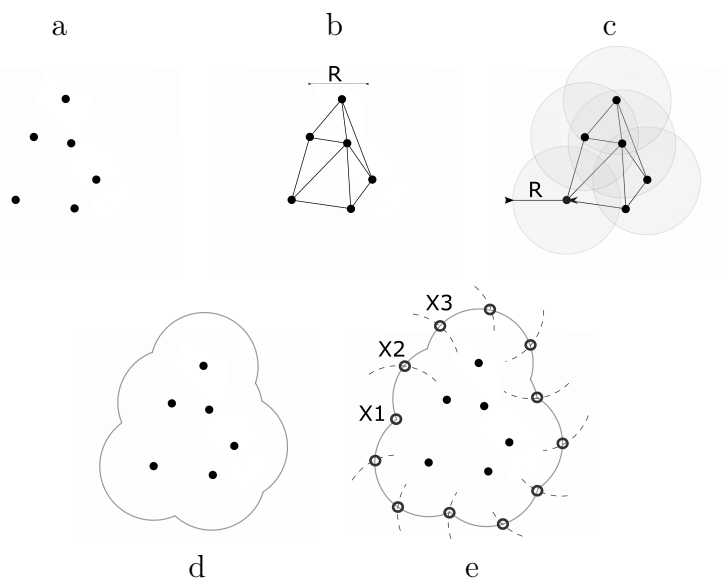


Fig. 3. One iteration of an infinite net construction; a) the given set of AS; b) Delaunay triangulation; c,d) surrounding equidistant construction; e) placing new layer of virtual nodes of the infinite net on the equidistant

The weighted average among the estimations under all the encountered conditions is used for the following final expression:

$$N \approx \sum_{j=1}^W \frac{T(\mathcal{A}^j)}{\mathbb{T}} N_j. \quad (1)$$

Each N_j is approximated as the total “successful” (animal at AS) observation time $T^*(\mathcal{A}^j)$ (summed over all the animals and all the cameras) divided by the “unit” time $T^0(\mathcal{A}^j)$ one average animal of the species under research spends at AS during the conditions \mathcal{A}^j

$$N_j \approx T^*(\mathcal{A}^j)/T^0(\mathcal{A}^j). \quad (2)$$

The value of $T^*(\mathcal{A}^j)$ may be measured from the observation survey

$$T^*(\mathcal{A}^j) = \sum_{i=1}^m T_i^*(\mathcal{A}^j), \quad (3)$$

where $T_i^*(\mathcal{A}^j)$ is the total “successful” observation time from the i -th camera. It must be noted that $T_i^*(\mathcal{A}^j)$ is summed up over all the animals independently, so if one elk spent 14 minutes at the i -th AS, then another one came and there were two elks during 7 minutes and then again only one remained for other 4 minutes, then $T_i^*(\mathcal{A}^j) = 14 + 7 + 7 + 4$ minutes.

For practical purposes, it is more convenient to use relative time in (2),

$$N_j \approx K(\mathcal{A}^j)/K^0(\mathcal{A}^j), \quad (4)$$

where $K(\mathcal{A}^j) = T^*(\mathcal{A}^j)/T(\mathcal{A}^j)$ and $K^0(\mathcal{A}^j) = T^0(\mathcal{A}^j)/T(\mathcal{A}^j)$.

While relative “successful” time $K(\mathcal{A}^j)$ can be measured directly from the survey, the “unit” relative “successful” time $K^0(\mathcal{A}^j)$ should be approximated a priori. Note that relying

on the Ergodic Hypothesis, $K^0(\mathcal{A}^j)$ may be regarded as the probability of one average animal to be seen at one average AS at a randomly selected moment $t_0 \in [0, T(\mathcal{A}^j)]$ during conditions \mathcal{A}^j .

Rewriting this probability using the Bayes naive classifier [38] with the assumption that for each fixed j the states $\alpha_k^j, k \in \overline{1, \Theta}$ are pairwise independent, I get

$$K^0(\mathcal{A}^j) = K^0 \frac{K^0(\alpha_1^j) \dots K^0(\alpha_\Theta^j)}{F(\alpha_1^j) \dots F(\alpha_\Theta^j)},$$

where

- K^0 is the probability to see an animal of the given species independently from the observation conditions; in practice it is the total time one animal of the given species spends at AS in any conditions divided by the total time of the observation surveys where that species was intended to see;
- $K^0(\alpha_k^j)$ is the probability to see an animal of the given species at AS during the state α_k^j ; in the context of computation, it is the total time of “animal at AS” during the state α_k^j divided by the total time devoted to the observation of the corresponding species during this state;
- $F(\alpha_k^j)$ is the frequency of the state α_k^j , which may be expressed as the total time of observation during the state α_k^j divided by the overall time of observation.

The resulting formula for the abundance estimate may be written as follows

$$N \approx \sum_{j=1}^W \frac{T(\mathcal{A}^j)}{\mathbb{T}} \frac{K(\mathcal{A}^j) F(\alpha_1^j) \dots F(\alpha_\Theta^j)}{K^0 K^0(\alpha_1^j) \dots K^0(\alpha_\Theta^j)}. \quad (5)$$

Values K^0 , $K^0(\alpha_k^j)$ and $F(\alpha_k^j)$ may be estimated by means of a special database (see Section 4) aggregating large amounts of information about the relative time one animal of the selected species spends at AS in different conditions. Now, let us proceed to the discussion of how this time may be computed.

3. Presence Time Estimation

Although the estimation of the time one animal spends at AS is definitely not an easy problem, it seems more tractable than the abundance estimation problem itself. The main benefit of the proposed Bayes approach (5) is that the more heterogeneous data is aggregated in the DB, the more exact abundance estimate is available in a wide variety of possible conditions. In this section I show how one can obtain the experimental information to improve the DB with the results of his local observations.

Within this section I suppose that the survey conditions and the duration of observation are the same for all the AS, so $\mathcal{A}^j = \mathcal{A}$ and $T(\mathcal{A}^j) = T(\mathcal{A}) = \mathbb{T}$.

Several techniques may be proposed for the measurement of the time one average unidentifiable animal spends at AS. The first and the most obvious is to identify the visitor, for example, using a GPS-collar [39], so the position of the animal is known exactly. If there are m' cameras and N' GPS-collars, then, for the k -th animal with a collar, the following approximation for the “unit” relative “successful time” holds:

$$\tilde{K}^0(k, \mathcal{A}) \triangleq \sum_{i=1}^{m'} \frac{\tilde{T}_i^*(k, \mathcal{A})}{T(\mathcal{A})}, \quad (6)$$

where $\tilde{T}_i^*(k, \mathcal{A})$ is the time that the k -th animal spent at the i -th AS during the conditions \mathcal{A} .

The resulting $K^0(\mathcal{A})$ is expressed as the average of the relative “successful” time among all the tracked animals

$$K^0(\mathcal{A}) = \frac{1}{N'} \sum_{k=1}^{N'} \tilde{K}^0(k, \mathcal{A}).$$

The same technique works if I mark several animals in some way (see e.g. [40]) to be able to recognize them and to measure $\tilde{T}_i^*(k, \mathcal{A})$ for each of them directly at the AS by camera-traps.

The third – a bit more complicated – approach is not connected with individual identification. Instead, the number of animals N' which visit the involved m' attractive spots is supposed to be known: either I take a semi-wild closed group with a priori known number of individuals or there exist some trusted methods to estimate the number of distinct animals visiting the AS (e.g. microsatellite analysis). Finally, I use the same logic as in (2) and (4), remembering that $\mathcal{A}^j = \mathcal{A}$ and $T(\mathcal{A}^j) = T(\mathcal{A}) = \mathbb{T}$:

$$N' = \sum_{i=1}^{m'} \frac{T_i^*(\mathcal{A})}{T^0(\mathcal{A})} = \sum_{i=1}^{m'} \frac{T_i^*(\mathcal{A})/\mathbb{T}}{K^0(\mathcal{A})}, \quad (7)$$

whence

$$K^0(\mathcal{A}) = \frac{T^*(\mathcal{A})}{N'\mathbb{T}}. \quad (8)$$

An interesting peculiarity of the last approach is that even the “hiders” (the animals that never appear before the cameras) influent the $K^0(\mathcal{A})$ obtained from (8) and through it the resulting estimation (5). Which means that the further surveys based on the combination of (5) and (8) are able to consider even the animals that do not visit the attractive spots.

The repeating presence time estimation surveys conducting in the (conventionally) same conditions will allow to express the expected bias in $K^0(\mathcal{A})$ for different \mathcal{A} and thereby the expected bias in the desired abundance estimation (5).

4. Database

The surveys database is the central but the most controversial element in the proposed method. As a nonspecialist in biology, ecology and nature conservation I might only hope to attract the attention of the specialists to the discussion around the structure and rules of filling of this database. The DB (see Fig. 4) could contain five types of columns:

- 1) species *id* (e.g. *Vulpes*);
- 2) mandatory survey properties $A = (\alpha_1, \dots, \alpha_\Psi)$ (e.g. day time, season, weather, climatic zone and so on); the results of the presence time estimation experiment (see Section 3) may be added to the DB only if all the mandatory survey properties are registered and the corresponding columns of the DB are filled;
- 3) optional survey properties $B = (\alpha_{\Psi+1}, \dots, \alpha_\Omega)$ (e.g. conditions of specific food reserves, distance from specific ecological zones and so on);
- 4) duration $T(\mathcal{A})$ of observation under the particular conditions $\mathcal{A} = (\alpha_1^j, \dots, \alpha_\Theta^j)$ (where \mathcal{A} covers all the properties from A and maybe some from B);

species	α_1	...	α_Ψ	...	α_Ω	$T(\mathcal{A})$	$K^0(\mathcal{A})$
...

Fig. 4. General form of the DB

5) corresponding average relative time $K^0(\mathcal{A})$ an individual animal spends at AS obtained from the presence time estimation experiment.

There are three main reasons to interact with the DB: read, write and enhance. Let us consider each of them briefly.

4.1. Read

The proposed example DB structure gives all the necessary information for applying (5) in an abundance estimation survey conducted for the species Sp :

- K^0 is approximated by the average value of column “ $K^0(\mathcal{A})$ ” among all the rows with the column “species” equal to Sp ;
- $K^0(\alpha_k^j)$ is approximated by the average value of “ $K^0(\mathcal{A})$ ” among all the rows with “species” equal to Sp and “ α_k ” column equal to α_k^j ;
- $F(\alpha_k^j)$ is approximated by the summary observation time “ $T(\mathcal{A})$ ” over all the rows with “ α_k ” equal to α_k^j divided by the total observation time “ $T(\mathcal{A})$ ” summed over the whole DB.

4.2. Write

To write operation is the simplest one to explain: the researcher conducts one of the presence time estimation experiments described in Section 3 and appends one new line to the DB, filling the columns according to the survey’s conditions and results.

As it was mentioned before, properties $A = (\alpha_1, \dots, \alpha_\Psi)$ must be measured and filled, while properties $B = (\alpha_{\Psi+1}, \dots, \alpha_\Omega)$ are optional.

4.3. Enhance

The enhancement of the DB is the responsibility of the community. This subsection gives just an example of a possible protocol. There may be two main roles of the DB users: researcher and expert. A researcher may propose: 1) new optional properties; 2) candidate optional properties to become mandatory; 3) to degrade some mandatory properties to optional; 4) to remove optional properties from the DB. A researcher may also vote for all the propositions made by the other researchers.

An expert is a researcher who decides the destiny of the properties according to the public opinion and statistics. For example, if the variation of some property (with all the other properties averaged) does not result in a significant variation of $K^0(\mathcal{A})$, then this property is a candidate for degrading from the mandatory or removing from the optional properties. And vice versa, if $K^0(\mathcal{A})$ varies significantly depending on a particular optional property, then this property becomes a candidate for mandatory.

There are a lot of statistical methods to evaluate the significance of dependency and the decision threshold [41–43]. This topic is not discussed here.

Conclusion

The proposed method's sketch is based on automated video-capturing at a moderate number easy to reach spots, so in the long term many labourious census works may be conducted easier, cheaper and cause less disturbance for the wild life. Information post-processing is strictly formalized, which leaves little chance for subjective alterations.

However, the method heavily relies on the volume and quality of the DB, which in its turn heavily relies on the efforts of the community. Public non-commercial solutions of complex problems (crowdsourcing) raises its popularity in such areas as software production (Freeware), investment (crowdfunding), scientific research (crowdsolving, civilian science). A DB similar to the proposed one, created by the community of zoologists and environment specialists, could not only benefit censuses, but also many ecological and behavioural studies.

The author is open for collaboration, both aimed at the practical adaptation of the proposed method for specific regions and at the development of the census methods theoretically.

Acknowledgements. *The author wants to thank Dr. Nikolay Markov (Institute of Plants and Animals, Ural Branch of Russian Academy of Sciences) for the fruitful discussions and critics and Evgeny Larin (Nature Park "Kondinskie Lakes") for practical insights.*

References

1. Borchers D.L., Buckland S.T., Zucchini W. *Estimating Animal Abundance: Closed Populations. Statistics for Biology and Health*. London, Springer, 2002. DOI: 10.1007/978-1-4471-3708-5
2. Seber G.A.F. *The Estimation of Animal Abundance and Related Parameters*. London, Macmillan Publishing Company, 1982.
3. Evans M.A., Bonett D.G., McDonald L. A General Theory for Analyzing Capture-Recapture Data in Closed Populations. *Biometrics*, 1994, no. 50, pp. 396–405. DOI: 10.2307/2533383
4. Arnason A.N., Schwarz C.J., Gerrard J.M. Estimating Closed Population Size and Number of Marked Animals from Sighting Data. *Journal of Wildlife Management*, 1991, no. 55, pp. 718–730. DOI: 10.2307/3809524
5. Karanth K. Estimating Tiger (*Panthera Tigris*) Populations from Cameratrapp Data Using Capture-Recapture Models. *Biological Conservation*, 1995, no. 71, pp. 333–338. DOI: 10.1016/0006-3207(94)00057-W
6. Carbone C., Christie S., Conforti K., Coulson T., et al. The Use of Photographic Rates to Estimate Densities of Tigers and Other Cryptic Mammals. *Animal Conservation*, 2001, no. 4, pp. 75–79. DOI: 10.1017/S1367943001001081
7. Cuellar E., Maffei L., Arispe R., Noss A. Geoffroy's Cats at the Northern Limit of Their Range: Activity Patterns and Density Estimates from Camera Trapping in Bolivian Dry Forests. *Studies on Neotropical Fauna and Environment*, 2006, no. 41, pp. 169–177. DOI: 10.1080/01650520600840001
8. Scott C.S., Linde E.T.O., Laura K.M., et al. The Use of Camera Traps for Estimating Jaguar *Panthera Onca* Abundance and Density Using Capture/Recapture Analysis. *Oryx*, 2004, vol. 38, no. 2, pp. 148–154.

9. Heilbrun R.D., Silvy L.J., Peterson M.J., Tewes M.E. Estimating Bobcat Abundance Using Automatically Triggered Cameras. *Wildlife Society Bulletin*, 2006, no. 34, pp. 69–73. DOI: 10.2193/0091-7648(2006)34[69:EBAUAT]2.0.CO;2
10. Maffei L., Noss A.J. How Small is Too Small? Camera Trap Survey Areas and Density Estimates for Ocelots in the Bolivian Chaco. *Biotropica*, 2008, no. 40, pp. 71–75.
11. Oliveira-Santos L.G.R., Zucco C.A., Antunes P.C., Crawshaw J.P.G. Is it Possible to Individually Identify Mammals with no Natural Markings Using Camera-Traps. *Mammalian Biology*, 2010, no. 75, pp. 375–378. DOI: 10.1016/j.mambio.2009.08.005
12. O'Connell A.F., Nichols J.D., Karanth K.U. *Camera Traps in Animal Ecology: Methods and Analyses*. N.Y., Springer, 2011. DOI: 10.1007/978-4-431-99495-4
13. Borchers D., Distiller G., Foster R., Harmsen B., Milazzo L. Continuous Time Spatially Explicit Capture Recapture Models, with an Application to a Jaguar Camera Trap Survey. *Methods in Ecology and Evolution*, 2014, no. 5, pp. 656–665. DOI: 10.1111/2041-210X.12196
14. Kohn M.H., York E.C., Kamradt D.A., Haught G., Sauvajot R.M., Wayne R.K. Estimating Population Size by Genotyping Faeces. *Proceedings of the Royal Society of London B: Biological Sciences*, 1999, no. 266, pp. 657–663. DOI: 10.1098/rspb.1999.0686
15. Lukacs P.M., Burnham K.P. Estimating Population Size from Dna-Based Closed Capture-Recapture Data Incorporating Genotyping Error. *Journal of Wildlife Management*, 2005, no. 69, pp. 396–403. DOI: 10.2193/0022-541X(2005)069<0396:EPSFDC>2.0.CO;2
16. Caniglia R., Fabbri E., Cubaynes S., Gimenez O., Lebreton J.D., Randi E. An Improved Procedure to Estimate Wolf Abundance Using Non-Invasive Genetic Sampling and Capture-Recapture Mixture Models. *Conservation Genetics*, 2012, no. 13, pp. 53–64. DOI: 10.1007/s10592-011-0266-1
17. Brinkman T.J., Person D.K., Smith W., Chapin F.S., et al. Using DNA to Test the Utility of Pellet-Group Counts as an Index of Deer Counts. *Wildlife Society Bulletin*, 2013, no. 37, pp. 444–450. DOI: 10.1002/wsb.270
18. Rodgers T.W., Janecka J.E. Applications and Techniques for Non-Invasive Faecal Genetics Research in Felid Conservation. *European Journal of Wildlife Research*, 2013, no. 59, pp. 1–16. DOI: 10.1007/s10344-012-0675-6
19. Schnell I.B., Sollmann R., Calvignac-Spencer S., Siddall M.E., Yu D.W., Wilting A., et al. iDNA from Terrestrial Haematophagous Leeches as a Wildlife Surveying and Monitoring Tool: Prospects, Pitfalls and Avenues to be Developed. *Frontiers in Zoology*, 2015, no. 12, pp. 1–14. DOI: 10.1186/s12983-015-0115-z
20. Fewster R.M. Some Applications of Genetics in Statistical Ecology. *ASTA Advances in Statistical Analysis*, 2017, vol. 101, no. 4, pp. 349–379. DOI: 10.1007/s10182-016-0273-0
21. Kery M. Estimating Abundance From Bird Counts: Binomial Mixture Models Uncover Complex Covariate Relationships. *The Auk*, 2008, vol. 125, no. 2, pp. 336–345. DOI: 10.1525/auk.2008.06185
22. Dail D., Madsen L. Models for Estimating Abundance from Repeated Counts of an Open Metapopulation. *Biometrics*, 2011, vol. 67, no. 2, pp. 577–587. DOI: 10.1111/j.1541-0420.2010.01465.x
23. Royle J.A., Chandler R.B., Sollmann R., Gardner B. *Spatial Capture-Recapture*. Amsterdam, Academic Press, 2014.
24. Ramsey D.S.L., Caley P.A., Robley A. Estimating Population Density from Presence-Absence Data Using a Spatially Explicit Model. *The Journal of Wildlife Management*, 2015, vol. 79, no. 3, pp. 491–499. DOI: 10.1002/jwmg.851

25. Buckland S.T., Thomas L. *Advanced Distance Sampling*. Oxford, Oxford University Press, 2007.
26. Buckland S.T., Anderson D.R., Burnham KP, Laake J.L. *Distance Sampling: Estimating Abundance of Biological Populations*. Springer Netherlands, 2012.
27. Buckland S.T., Rexstad E.A., Marques T.A., Oedekoven C.S. *Distance Sampling: Methods and Applications*. N.Y., Springer, 2015. DOI: 10.1007/978-3-319-19219-2
28. Royle J.A., Nichols J.D. Estimating Abundance from Repeated Presence Absence Data or Point Counts. *Ecology*, 2003, no. 84, pp. 777–790. DOI: 10.1890/0012-9658(2003)084[0777:EAFRPA]2.0.CO;2
29. Hutchinson J.M.C., Waser P.M. Use, Misuse and Extensions of 'Ideal Gas' Models of Animal Encounter. *Biological Reviews of the Cambridge Philosophical Society*, 2007, no. 82, pp. 335–359. DOI: 10.1111/j.1469-185X.2007.00014.x
30. Rowcliffe J.M., Field J., Turvey S.T., Carbone C. Estimating Animal Density Using Camera Traps Without the Need for Individual Recognition. *Journal of Applied Ecology*, 2008, no. 45, pp. 1228–1236. DOI: 10.1111/j.1365-2664.2008.01473.x
31. Foster R.J., Harmsen B.J. A Critique of Density Estimation from Cameratrapp Data. *The Journal of Wildlife Management*, 2012, no. 76, pp. 224–236. DOI: 10.1002/jwmg.275
32. Lucas T.C.D., Moorcroft E.A., Freeman R., Rowcliffe J.M., Jones K.E. A Generalised Random Encounter Model for Estimating Animal Density with Remote Sensor Data. *Methods in Ecology and Evolution*, 2015, no. 6, pp. 500–509. DOI: 10.1111/2041-210X.12346
33. Ferguson P.F.B., Conroy M.J., Hepinstall-Cymerman J. Occupancy Models for Data with False Positive and False Negative Errors and Heterogeneity Across Sites and Surveys. *Methods in Ecology and Evolution*, 1994, no. 6, pp. 1395–1406.
34. McCarthy M.A., Moore J.L., Morris W.K., Parris K.M., Garrard G.E., Vesk P.A., et al. The Influence of Abundance on Detectability. *Oikos*, 2013, no. 122, pp. 717–726. DOI: 10.1111/j.1600-0706.2012.20781.x
35. Voronoi G. Nouvelles applications des parametres continus a la theorie des formes quadratiques. *Journal für die Reine und Angewandte Mathematik*, 1908, no. 133, pp. 97–178. (in Italian)
36. Aurenhammer F., Klein R., Lee D.T. *Voronoi Diagrams and Delaunay Triangulations*. Singapore, World Scientific Publishing Company, 2013. DOI: 10.1142/8685
37. Delaunay B.N. Neue Darstellung der geometrischen Kristallographie. *Zeitschrift für Kristallographie*, 1932, no. 84, pp. 109–149. (in German)
38. Hand D.J., Yu K. Idiot's Bayes – not so Stupid after All? *International Statistical Review*, 2001, no. 69, pp. 385–398.
39. Hulbert I.A.R., French J. The Accuracy of GPS for Wildlife Telemetry and Habitat Mapping. *Journal of Applied Ecology*, 2001, no. 28, pp. 869–878. DOI: 10.1046/j.1365-2664.2001.00624.x
40. Braun C. *Techniques for Wildlife Investigation and Management*. Wildlife Society, 2005.
41. Clark J.S. *Models for Ecological Data: An Introduction*. Princeton, Princeton University Press, 2007.
42. Watt T.A., McCleery R.H., Hart T. *Introduction to Statistics for Biology*. London, Chapman and Hall/CRC, 2007.
43. Wackerly D., Mendenhall W., Scheaffer R. *Mathematical Statistics with Applications*. Thomson Brooks/Cole, 2008.

Received August 8, 2018

ИСПОЛЬЗОВАНИЕ БОЛЬШИХ ДАННЫХ В ОЦЕНКЕ ЧИСЛЕННОСТИ ИНДИВИДУАЛЬНО НЕИДЕНТИФИЦИРУЕМЫХ ЖИВОТНЫХ С ПОМОЩЬЮ КАМЕР-ЛОВУШЕК НА ПРИВЛЕКАТЕЛЬНЫХ СТРАТАХ

Е.Е. Иванко, Институт математики и механики УрО РАН, г. Екатеринбург, Российская Федерация; Уральский федеральный университет, г. Екатеринбург, Российская Федерация

Камеры-ловушки являются относительно новым, но уже популярным инструментом оценки численности индивидуально неидентифицируемых животных. Хотя камеры-ловушки удобны, при их применении остается как ряд теоретических трудностей (таких как пространственная автокорреляция или проблема ложноотрицательных наблюдений), так и чисто практические сложности, связанные, например, с трудоемкостью сбора рандомизированных данных. В данной статье автор предлагает альтернативный метод организации учета, позволяющий избежать указанных проблем. Предложенный подход основан на сборе видеоматериала с помощью камер-ловушек, расположенных в местах естественного притяжения животных. По собранным видеоданным рассчитывается частота посещений, преобразуемая далее в оценку численности животных в исследуемой области. Ключом к такому преобразованию служат корректирующие коэффициенты, вычисляемые для совокупности конкретных условий наблюдения с помощью применения Байесовского классификатора к масштабной базе данных наблюдений в различных условиях. В долгосрочной перспективе предложенный подход позволит проводить трудоемкие работы по оценке численности индивидуально неидентифицируемых животных легче и дешевле, а также приводить к меньшим вмешательствам в дикую среду обитания. Обработка полученных видеоданных строго формализована, так что предмета для субъективных разногласий при учете практически не остается. Изложенный в работе метод существенно зависит от объема и качества базы данных наблюдений, которая, в свою очередь, существенно зависит от усилий заинтересованного сообщества. Хотя конструирование подобной базы данных может быть сложной и противоречивой задачей, ее решение представляется существенно более легким, чем решение исходной задачи оценки численности отдельно для каждого конкретного случая. Создание подобной базы поможет не только при учете численности неидентифицируемых животных, но также обеспечит богатый источник данных для различных поведенческих исследований.

Ключевые слова: оценка численности; камеры-ловушки; большие данные; наивный Байесовский классификатор.

Евгений Евгеньевич Иванко, доктор физико-математических наук, заведующий лабораторией анализа сложных систем отдела вычислительных систем, Институт математики и механики УрО РАН (г. Екатеринбург, Российская Федерация); Уральский федеральный университет (г. Екатеринбург, Российская Федерация), evgeny.ivanko@gmail.com.

Поступила в редакцию 8 августа 2018 г.